Geographical information systems and tropical diseases

Stan Openshaw Centre for Computational Geography, School of Geography, University of Leeds, Leeds, LS2 9JT, UK

Keywords: geographical information systems, tropical diseases

The geographical information system (GIS) revolution started in the mid 1980s and has since spread to all countries of the world. A GIS is a computer software package that is concerned with the capture, storage, manipulation, query, display and analysis of any and all types of geographical information. (BURROUGH, 1986; CHORLEY, 1987; LAURIN & THOMPSON, 1992; ESRI, 1993). There are now many different GIS packages available throughout the world. The unique attraction of GIS is that it adds a geographical dimension to information held in computer systems that previously lacked it, or was stored using database software which was inadequate to handle and manipulate the geography associated with the data.

Traditional geographical information consists of virtually everything shown on maps of any kind; for example, at large scale topographical maps, geographically underground pipes and wires, town plans, and thematic maps such as those showing land-use, geology, soils, climatic variation, and disease. It had been known since the 1970s how to store these cartographic map objects in a digital form; essentially as very large numbers of two-dimensional map co-ordinates (longitude and latitude) that represent the elemental points, lines, and polygons that maps are built from and can be decomposed into. The GIS revolution occurred when suddenly, over a short period of time, software availability and dramatic price and performance improvements in computer hardware made it feasible and economic to start converting all these traditional geographical information sources stored in analogue form on paper maps into a digital format suitable for computer processing. In many countries the first task was to regenerate the existing national paper map resources in a digital form so that the cartography could be done automatically. So, at its simplest, GIS is little more than a spatial database and mapping system that runs on a workstation of some kind and which greatly eases the task of drawing maps. This was undoubtedly convenient, it greatly speeded the mapping process and made computer storage and retrieval of paper map records far more efficient. However, there is much more to GIS than its origins in the automation of cartography and in land and facilities management might indicate. Subsequent developments permit geographical references to be added to many other databases, where previously they were present only in an implicit and indirect form; for example, as a postal address. Furthermore, the availability of digital map data in a common format greatly facilitates the task of data integration, cross referencing and, perhaps more relevant in the present context, modelling and analysis of the spatial patterns and processes locked up in the data.

Today, geographical information extends far beyond the traditional map-based sources but includes all types and forms of information that can be related to a map. These extend from databases about people to remote-sensing systems based on space platforms, and include real-time tracking and other automatic data capture devices in which spatial location is an important feature. Properly developed and applied by competent and modern organizations, GIS provides an integrating computer database technology that allows all types of geographical information to be integrated, interrelated, and handled in a unified way. The task now is to start to exploit the tremendous explosion in the availability of spatial information that has occurred during the last 5 years as a result of GIS and also, more generally, because of the increased use of information technology in all areas of life and in all parts of the world. There has never been so much information available about the world in which we live. The challenge is to convert the dreams of the GIS missionaries and the sometimes exaggerated sales talk of the GIS vendors into useful, valuable, and life-critical systems and applications relevant to tropical disease management, prevention, and research.

There are many areas of medical interest that could well benefit from the availability of GIS and the increasingly spatial data-rich world in which we live. Some of the current principal impediments are discussed below. (i) Organizational problems, when the impact of GIS is greatly diminished and sometimes totally negated, at least, in the short term, due to a failure to realize or accept the corporate nature of both the technology and the associated databases with the consequential need for management restructuring. (ii) Problems of access to the necessary information, either because of the cost of acquiring copyright information or because for various reasons the critical information sources are either not yet available or are not in a suitable form for analysis. (iii) Lack of attention by the GIS vendors to the development of applicable spatial analysis and modelling technologies, as the principal markets have so far been mainly in areas related to the capture, management, and mapping of geographical information and not its analysis. (iv) Absence of a strong analytical tradition in areas where previously analysis was either impossible or impracticable, which continues to hinder future potential applications that are largely analysis-based. (v) Negative or unnecessarily defensive attitudes that allow largely unwarranted fears of a breach of data confidentiality or theoretical concerns about the misuse of personal information to prevent any analysis being performed regardless of the public good—an argument that should ideally override such concerns, particularly in a medical context. (vi) Lack of demonstrated benefit, which makes the first application so difficult to establish and often engenders a high level of exaggeration and unrealistic expectations as innovators struggle to develop applications against well entrenched opposition, ignorance and lack of resources.

Let me be quite frank. GIS can save lives but its use does require a degree of faith and the adoption of something longer than a short-term view. GIS, is or soon will be, an essential component for any and all medical computer systems in virtually all contexts throughout the world. Before long it will be unthinkable not to use GIS as a management and analytical support tool; for example, to monitor disease databases for patternning, to ensure evenness in the provision of medical facilities, and to watch the performance of care delivery systems. So whatever 'sorry' the GIS vendor may use to persuade you to buy a GIS today, rest assured that (i) it largely does not seem to matter much which one you buy as long as it is a mature and well established product with adequate customer support; (ii) most of the benefits vendors describe will be achievable, but not always without considerable effort, and the biggest returns may well be delayed until the next generation of system is purchased, so portability of the data infrastructure is critical; and (iii) the GIS will probably not do much to justify its existence.
this year or next, and should be regarded as a longer term investment.

So what benefits might GIS be expected to offer to tropical medicine that are sufficient to justify the expenditure of scarce resources? Currently there is not much demonstrated practical evidence that can be used to convince the doubters. A strong element of faith and optimism is initially required but, given good will and a little luck, GIS may well be expected to offer at least some of the following benefits in a tropical medicine context.

(i) The most immediate use of GIS is an enhanced ability to map spatial information relating to the incidence of one or more diseases that seem to be involved; for instance malarial infection and water bodies. This type of geographical association by map overlay is easily performed and might well be extremely useful where the disease patterns are fairly well defined. Similarly, it is fairly easy to apply spatial analysis based on low technology graphic methods to tease out interesting patterns of incidence and relationships (see, for example, Haslett et al., 1991 and Bailey & Gatrell, 1995). However, geographical analysis will never prove or establish causation, only suggest, stimulate and inform: maybe that is more than sufficient.

(ii) More complex patterns require more sophisticated technology. Exploratory spatial data analysis of epidemiological databases is potentially extremely useful in dealing with less obvious patterns; for instance, where there is latency before symptoms appear. Indeed, once a database has geography attached to it there is both an opportunity and an imperative to perform various types of exploratory geographical analysis. The problem here is that GIS is quite good at answering precisely formulated questions that either reflect knowledge or hypotheses. For example, what is the malaria rate within 2 km of all lakes, or within 1 km of the river? One then determines whether or not the answers are in any way statistically unusual or significant, or otherwise worthy of further investigation. Current GISs do not offer much assistance here. Another potentially much more serious problem arises if one does not know what questions to ask or what spatial queries to make. There is as yet no magic 'intelligent query system' that will scan a database and suggest queries that might be worth investigating; although there should be! If one has no a priori knowledge-based question to ask, or if the patterns or geographical associations are too complex to spot by looking at multicoloured maps, then he or she is stuck! Ironically, the better the quality and the resolution of the database, the more spatial information that is available, the more disease types or medical variables that are included, the worse these problems of exploratory analysis become. The GIS is therefore basically a first step in creating a database resource that can subsequently be examined by methods that are currently not included in many (or any) of the proprietary GIS systems. These problems are not criticisms of GIS, rather a commentary on what is missing from currently available commercial systems. Given time, the necessary technology will appear—either for use within GISs or as 'stand-alone' systems that can be fed with GIS data. The question is, can one afford to wait?

(iii) Currently several research tools exist that could be used to perform more sophisticated exploratory spatial analysis of medical data. The geographical analysis machine (GAM/1) of Openshaw et al. (1987) and the artificial intelligence (AI)-based pattern-seeking creatures of Openshaw (1994, 1995) all offer an exploratory pattern hunting capability. If there is an interest in spatial hypothesis testing, then the methods of Diggle & Chetwynd (1991) and Diggle et al. (in press) might be worth considering. If the search is for geographical associations, then the geographical correlates exploration machine (GCEM) of Openshaw et al. (1990) offers a means of scanning permutations of map coversages for unusual associations with disease patterns. However, it is only fairly recently that academic researchers have begun to realize the importance of developing exploratory tools that ordinary end-users who are not skilled statisticians can use; see, for example, Openshaw & People (1996) who describe a mixture of AI technology and computer animation to develop an intuitively obvious approach to spatial analysis. The users discover what is going on by watching an animation of a spatial analysis tool operating on their data. If patterns are present they will slowly emerge.

(iv) A different set of applications will involve the spatial modelling of the functioning of medical systems so that their territorial patterning might be optimized. What is the best location for a new hospital? What
changes should be made to an existing network of facilities to optimize accessibility levels? Current GISs already offer fairly basic spatial decision support technologies, although really important applications tend to have highly customized software systems; for example, the so-called 'intelligent' GIS of BIRKIN et al. (1996). Again, the technology is known and fairly well developed in retail contexts; however, the potential public benefits from medical applications are just as real but are at present grossly underdeveloped.

(vi) Other, more fanciful, applications of GIS may well involve spatial process modelling as the basis for clever detective work. There is often a strong geographical effect in many tropical disease distributions. For instance, the distance travelled by female mosquitoes from a source of water. Computer models can be run in reverse from disease incidence to identify source areas. Again GIS provides the necessary data infrastructure and in this case maybe also some of the modelling technology needed for the subsequent detective work.

(vii) In the near future there will be increasingly important real-time applications of GIS as a means of tracking the diffusion of diseases in time and space. As the ability to track movement in space and time via automated monitoring and data capture devices improves, so GIS offers the necessary data management technology needed for real-time analysis and decision-making. Maybe a little imagination is a key ingredient here.

In conclusion, it can be argued that GIS is an increasingly essential technology. It already offers an integrating computer system for providing a total data view. As improved exploratory data analysis and spatial modelling tools are developed, so the capital investment in GIS will start to return substantial medical benefits. Initially, getting it all going is a matter of faith, although ultimate success is virtually always assured. The only uncertainty is when and not if.

References